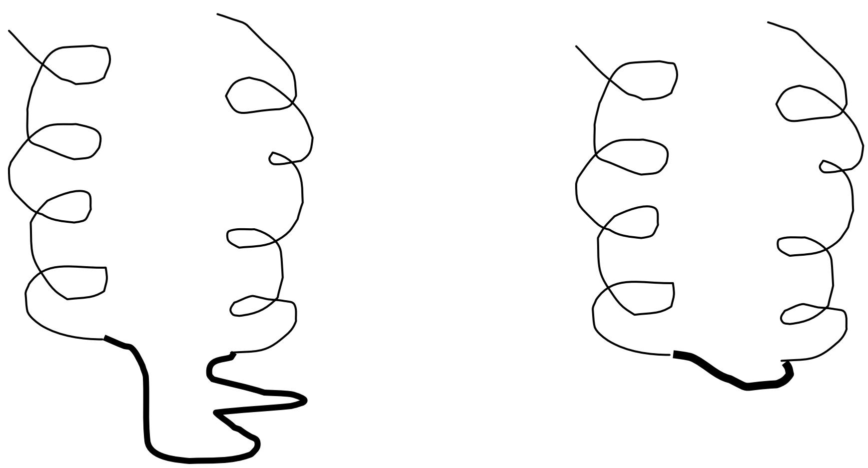
Biological motivation for Lagrange multipliers and rotation matrices
Structural Overlap

We wish to have a quantitative measure to determine if the shapes of two proteins are similar. This is important for many tasks. High degree of similarity between protein shapes suggests an evolutionary link and will make it possible for us to seek and study protein ancestors and directions of evolution. High degree of similarity may also suggest similar functions of the structurally similar proteins. If we know the function of only a subset of the protein family it is common to assume a similar function for other proteins of the same family with similar structures.

Yet another reason for the determination of structural similarity is the use of structural families to create detailed atomic models. We can use a family of similar proteins to create a structural alignment of this family. A structural alignment tells us which parts of the structure are well preserved during the evolution and therefore can be effectively modeled for the new protein. An example for two structures of the same family that share only fragments of structures is schematically drawn below



We should be able to detect the similarity between the fragments (helices), and point to the loop as the structural segment that deviates the most, regardless of the sequence. Since the identities of the amino acids are not used (only the C_α positions) in comparing backbone structures highly remote evolutionary connections may be observed. This is our final goal. We start however with the (much) simpler case of overlapping proteins of the same length (no alignment is necessary just proper measure of their distance).

The comparison of two protein structures is actually less trivial than it may seem from first sight. Before sinking into the derivation of the algorithm, we need to make two intermediate stops and to arm ourselves appropriately for the task at hand. We need a quick introduction to Lagrange's multipliers and rotation matrices.

Lagrange multipliers

Consider an optimization problem with a constraint. For example, we wish to minimize a function $f(x,y,z) = x^2 + y^2 + z^2$ with respect to the variables x, y, z. The global minimum is clearly when $(x = 0, y = 0, z = 0)$.

However we now add a constraint $\sigma = x + y + z - 1 = 0$. Hence our solution should be a minimum of f and at the same time satisfies the condition σ. In principle there can be more than one function that we wish to minimize and more than one constraint. However, the discussion here is limited to one function, but potentially more than one constraint. Note that each of the constraints removes one variable. For example the above constraint established a relation between $x, y,$ and z, and there are no more independent. This suggests the first approach we can use to solve the constrained problem.

How can we solve this problem? One approach is to use the constraint(s) to solve for one of the variables. For example, we can use $z = 1 - x - y$ and minimize the new function $f(x,y,z) \equiv \overline{f}(x,y) = x^2 + y^2 + (1 - x - y)^2$. The minimization in this case is trivial and can be done analytically (no need for MATLAB or numerical analysis). We have

$$\frac{d\overline{f}}{dx} = \frac{d\overline{f}}{dy} = 0 \qquad \frac{d\overline{f}}{dx} = 2x - 2(1 - x - y) = 0 \qquad \frac{d\overline{f}}{dy} = 2y - 2(1 - x - y) = 0$$

$$\Rightarrow x = 1/3 \quad y = 1/3 \quad \text{z=}1/3$$

The difficulty with the above (which is nevertheless an exact procedure) is that it is not always easy to solve explicitly for one of the variables. For example consider the constraint: $\sigma = \exp\left[x^2 + y^{\frac{1}{5}} + 1/z\right] + y^2 + z^7 - x = 0$, can you express z in terms of x and y? Moreover, the problem is treated in a very asymmetric way, who said that we should substitute z and not y or x? This may not seems important but in large systems such choices can lead to numerical instabilities.

The Lagrange multipliers approach makes it possible to do the constrained optimization in a symmetric way almost like an optimization without constraints. We consider a new function $g\left(x,y,z,\lambda\right)$ with one more variable, λ compared to three variables of $f\left(x,y,z\right)$. The new function is given by

$$g\left(x,y,z,\lambda\right)=f\left(x,y,z\right)+\lambda\sigma\left(x,y,z\right)$$

If we now wish to determine a stationary point of $g\left(x,y,z,\lambda\right)$, first derivatives are useful

$$\frac{dg}{dx}=\frac{df}{dx}+\lambda\frac{d\sigma}{dx}=0$$
$$\frac{dg}{dy}=\frac{df}{dy}+\lambda\frac{d\sigma}{dy}=0$$
$$\frac{dg}{dz}=\frac{df}{dz}+\lambda\frac{d\sigma}{dz}=0$$
$$\frac{dg}{d\lambda}=\sigma=0$$

Let us look at the concrete example we had above and solve it in the framework of Lagrange's multipliers

$$g\left(x,y,z,\lambda\right)=x^2+y^2+z^2+\lambda\left(x+y+z-1\right)$$
$$\frac{dg}{dx}=2x+\lambda=0$$
$$\frac{dg}{dy}=2y+\lambda=0$$
$$\frac{dg}{dz}=2z+\lambda=0$$
$$\frac{dg}{d\lambda}=(x+y+z-1)=0$$

The four equations with four unknown can be solved by eliminating λ from the first two equations –

$$\frac{dg}{dx}-\frac{dg}{dy}=2x-2y=0\Rightarrow x=y$$
$$\frac{dg}{dx}-\frac{dg}{dz}=2x-2z=0\Rightarrow x=z$$

and now substituting in the last equation to have $3x-1=0 \quad x=1/3$. We therefore also have $x=1/3 \;\; y=1/3 \;\; z=1/3 \;\; \lambda=-2/3$. It is not really necessary here to determine the Lagrange multiplier since it is only a tool to solve the constrained optimization problem.

However it is interesting to find out what it does. The Lagrange multiplier is multiplying a term which is the constraint derivative and balances the gradient of the original function. This is why the derivatives of the constraints are sometimes called "constraint force".

Lagrange's multipliers are extremely useful in many branches of mechanics. For example, in simulation of robotic motions, one usually assumes that the length of the robot's hand is fixed and only the orientation is changing when we optimize the structure of the robot. The constant length of the robot hand can be modeled with a constraint and the use of Lagrange's multipliers.

Of course rather than solving the constrained equations analytically we could search for a local minimum numerically, even if the expression for the gradient is not as simple as we had above. We will discuss minimization algorithm towards the end of the class. This minimization algorithm can be easily extended to the constrained minimization problem once the Lagrange's multipliers idea is introduced. Note however that g is not a minimum since the second derivative matrix is not (as required from a true minimum) positive definition. The only condition that we use above is that the gradient will be zero. As a result the problem of constrained minimization is usually a bit more complicated that the problem of unconstrained minimization is usually easier to solve than the problem of constrained minimization.